BMJ Open

Cohort profile: The Applied Public-Private Research enabling OsteoArthritis Clinical Headway (IMI-APPROACH) study: a 2-year, European, cohort study to describe, validate and predict phenotypes of osteoarthritis using clinical, imaging and biochemical markers

Eefje M van Helvoort ,[1] Willem E van Spil,[1] Mylène P Jansen ,[1] Paco M J Welsing,[1] Margreet Kloppenburg,[2,3] Marieke Loef ,[3] Francisco J Blanco ,[4] Ida K Haugen,[5] Francis Berenbaum,[6] Jaume Bacardit,[7] Christoph H Ladel,[8] John Loughlin,[9] Anne C Bay-Jensen,[10] Ali Mobasheri,[11] Jonathan Larkin,[12] Janneke Boere,[13] Harrie H Weinans,[1,14] Agnes Lalande,[15] Anne C A Marijnissen,[1] Floris P J G Lafeber[1]

To cite: van Helvoort EM, van Spil WE, Jansen MP, *et al*. Cohort profile: The Applied Public-Private Research enabling OsteoArthritis Clinical Headway (IMI-APPROACH) study: a 2-year, European, cohort study to describe, validate and predict phenotypes of osteoarthritis using clinical, imaging and biochemical markers. *BMJ Open* 2020;**10**:e035101. doi:10.1136/bmjopen-2019-035101

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2019-035101).

Received 18 October 2019
Revised 15 April 2020
Accepted 06 May 2020

For numbered affiliations see end of article.

Correspondence to
Dr Eefje M van Helvoort;
E.M.vanHelvoort-3@umcutrecht.nl

ABSTRACT

Purpose The Applied Public-Private Research enabling OsteoArthritis Clinical Headway (APPROACH) consortium intends to prospectively describe in detail, preselected patients with knee osteoarthritis (OA), using conventional and novel clinical, imaging, and biochemical markers, to support OA drug development.

Participants APPROACH is a prospective cohort study including 297 patients with tibiofemoral OA, according to the American College of Rheumatology classification criteria. Patients were (pre)selected from existing cohorts using machine learning models, developed on data from the CHECK cohort, to display a high likelihood of radiographic joint space width (JSW) loss and/or knee pain progression.

Findings to date Selection appeared logistically feasible and baseline characteristics of the cohort demonstrated an OA population with more severe disease: age 66.5 (SD 7.1) vs 68.1 (7.7) years, min-JSW 2.5 (1.3) vs 2.1 (1.0) mm and Knee injury and Osteoarthritis Outcome Score pain 31.3 (19.7) vs 17.7 (14.6), except for age, all: p<0.001, for selected versus excluded patients, respectively. Based on the selection model, this cohort has a predicted higher chance of progression.

Future plans Patients will visit the hospital again at 6, 12 and 24 months for physical examination, pain and general health questionnaires, collection of blood and urine, MRI scans, radiographs of knees and hands, CT scan of the knee, low radiation whole-body CT, HandScan, motion analysis and performance-based tests.

After two years, data will show whether those patients with the highest probabilities for progression experienced disease progression as compared to those wit lower probabilities (model validation) and whether phenotypes/endotypes can be identified and predicted to facilitate targeted drug therapy.

Trial registration number NCT03883568

Strengths and limitations of this study

- The Applied Public-Private Research enabling OsteoArthritis Clinical Headway (APPROACH) cohort is part of a larger consortium, bringing together a highly qualified and multidisciplinary group of stakeholders in the form of a public–private partnership of engaged, knowledgeable and complementary industrial, academics and patient experts.
- The APPROACH cohort is unique in its selection process, recruiting patients from existing cohorts based on machine learning models with encouraging results of which the actual utility needs to be demonstrated at the end of the 2-year follow-up.
- The APPROACH cohort will provide 2-year follow-up data of 297 knee osteoarthritis patients including conventional and novel, explorative, imaging, biochemical, clinical and demographic (bio)markers according to strict protocols for acquisition and evaluation with the aim to identify phenotypes and develop predictive models for progression of these phenotypes.
- The main limitations of the study are the descriptive phase in which the study is at present and the still limited number of 297 patients related to the large number of outcome parameters.

INTRODUCTION

Osteoarthritis (OA) is characterised by changes in all (peri)articular tissues,[1 2] causing pain, stiffness and loss of function, usually following a slowly progressive and nonlinear course.[2] OA of the knee is the most common and most disabling, accounting for 83% of total OA burden.[3] In 2010, the global prevalence of knee OA was estimated to be 4.7% in females and 2.6% in males and incidence peaked around the age of 50.[4] Knee OA accounted at that time for 14.218 of total years lived with disability. This is a 64.8% increase compared with 1990 (8.627), emphasising the increasing burden of OA.[3] Estimated healthcare costs of knee OA are €4.257 (€383–€7.675) per patient per year. Non-healthcare-related costs of knee OA, like productivity loss, are estimated to be €1.519 per patient per year.[5] Ageing of the population, an increasing active life style at older age, and the current obesity pandemic all contribute to an even further increase of the incidence and prevalence of OA and its societal burden.[6]

Despite this growing OA burden and the still unmet need for effective treatment, pharmaceutical companies seem to have lost their confidence in drug development because clinical trials with potential disease-modifying OA drugs (DMOADs) could not demonstrate efficacy. This disappointing result likely has multiple origins.

The typically slow and heterogeneous OA course makes trials easily fall short in terms of size and length for demonstrating treatment efficacy.[7] This issue is further aggravated by the use of relatively insensitive outcome measures (patient-reported outcome measurements), pain and radiographic joint space changes (X-ray), required by regulatory agencies for a drug to be certified as a DMOAD. Moreover, an incomplete understanding of the OA pathobiology obscures identification of proper treatment targets. This is complicated by the increasing knowledge that the pathobiological mechanisms driving the OA process differ between patients, (type of) joints and disease stages.[2]

This to-date concept of a highly heterogeneous disease contrasts with the one-size-fits-all treatment approach used in most trials and the focus on radiographic joint space narrowing (JSN) and pain as outcome parameters.

New (combinations of) sensitive and robust (bio) markers could importantly contribute to overcome the aforementioned challenges, improving the design of clinical trials in the OA field. Biomarkers with the ability to predict the likely disease course in an untreated individual, viz. prognostic markers, could be employed to identify subjects that will show significant progression of the disease on relevant outcome parameter(s) over the study period. Biomarkers that show a biological response to treatment, response markers, could serve as sensitive outcome parameters, supplementing (or even replacing) radiographic joint space changes and MRI read-outs. These biomarkers could also help to identify vital components of the OA pathobiology and with that distinguish between phenotypes/endotypes. This will help to forecast the potential response to treatments targeted to specific mechanisms. Altogether, such biomarkers could importantly improve the quality and effectiveness of trials of potential DMOADs and joint preserving surgical treatments, in terms of selection of study participants, outcome parameters, and study size and length.[8]

Applied Public-Private Research enabling OsteoArthritis Clinical Headway

Although currently available cohort studies, like the Dutch CHECK[9] and the US OAI with the FNIH[10] have increased our knowledge of the disease, these attempts still have not resulted in clearly distinctive phenotypes/endotypes with predictive biomarkers. Therefore, the current Applied Public-Private Research enabling OsteoArthritis Clinical Headway (APPROACH) cohort uses a novel strategy and extends on previous studies in several ways. The study is part of a larger consortium being conducted under the Innovative Medicine's Initiative, bringing together a highly qualified and multidisciplinary group of stakeholders in the form of a public–private partnership of engaged, knowledgeable and complementary industrial, academic, and patient experts.

The APPROACH cohort is unique in its attempt to recruit patients primarily from existing cohorts using machine learning (ML) models (adjusted to the specific cohorts) trained using patient data from the CHECK cohort to increase the likelihood of radiographic joint space width (JSW) loss and/or knee pain progression during a limited, 2-year follow-up period. The relative short 2-year period is chosen to facilitate translation of results to pragmatic trial design.

In addition to this unique preselection of patients, the APPROACH cohort combines a very broad spectrum of conventional and novel, explorative, imaging, biochemical, clinical and demographic markers. Using data science techniques suitable to analyse these 'big data', algorithms of biomarkers will identify and predict phenotypes/endotypes of OA that share distinct underlying pathobiological mechanisms with their structural and function consequences, relevant for practical and targeted clinical trials.

The objectives of the cohort study are (https://www.approachproject.eu):

- To validate and refine the prediction model for sustained pain and decrease in (minimum) JSW as developed for the selection of patients.
- To develop and validate sensitive markers of/predictive for OA progression by use of conventional and novel clinical, imaging, and biochemical (bio) markers.
- To discover and predict novel OA phenotypes (eg, post-traumatic, metabolic, ageing, inflammatory, bone driven and genetic) and (their) disease progression.
- To prospectively describe in detail the discovered phenotypes by use of conventional and novel clinical, imaging and biochemical (bio)markers.

COHORT DESCRIPTION

The prospective follow-up of the 297 included patients will be 2 years. A large spectrum of conventional and novel (bio)markers for discovery (baseline, 1-year and 2-year follow-up), and prediction (baseline and change over six months) of knee OA phenotypes/endotypes will be gathered.

Patient selection

Patients were stepwise selected for a high chance of structural progression (JSN) and/or pain progression/sustained severity over two years, using two ML models, for the likelihood of each patient to be a 'progressor'.

The selection process will be described in detail elsewhere.[11] In summary, patients with predominant tibiofemoral OA were selected from five European observational OA cohorts (CHECK,[9] HOSTAS,[12] MUST,[13] PROCOAC[14] and DIGICOD; for cohort details see online supplementary file 1) using an ML approach, trained on longitudinal data from the CHECK cohort, and adjusted for the specific cohorts using the available data from each of the cohorts. Separate models for prediction of structural progression and for sustained significant knee pain or pain progression were used. Structural progression was defined as a reduction in JSW of ≥0.3 mm per year over a period of 2–3 years (0.7 mm is the minimal detectable difference in radiographic JSW).[15] Sustained significant pain and pain increase were defined as at least one of the three following: Knee injury and Osteoarthritis Outcome Score (KOOS) pain (on a 0–100 scale) increase ≥5/year and ≥40 at two years, KOOS pain increase ≥10/year and ≥35 at two years or ≥40 at both baseline and two years. Three types of progression were defined: pain progression, structural progression, and both pain and structural progression.

All identified patients of the existing cohorts (ranking the highest for predicted progression in the first ML model) willing to participate were invited for a screening visit. During this visit, inclusion and exclusion criteria were checked and an index knee was selected based on American College of Rheumatology (ACR) criteria.[16] If both knees fulfilled the criteria, patients indicated their own index knee based on severity of complaints, in case equal the right knee was selected as the index knee. Key predictors from the first predictive ML model, for example, KOOS[17] and Knee Image Digital Analysis (KIDA) parameters,[18] were collected and fed into a subsequent predictive ML model that was uniform for all cohorts. The patients who ranked the highest in this second model were included and invited for a baseline visit.

Because the existing source cohorts could not all provide sufficient patients due to the selection process, patients withdrawing consent and non-compliance with inclusion criteria, a small number of additional patients were recruited from outpatient departments and invited for a screening visit (see figure 1). These newly recruited patients were also ranked and selected using the second, uniform predictive ML model.

Inclusion criteria

- Able to walk unassisted.

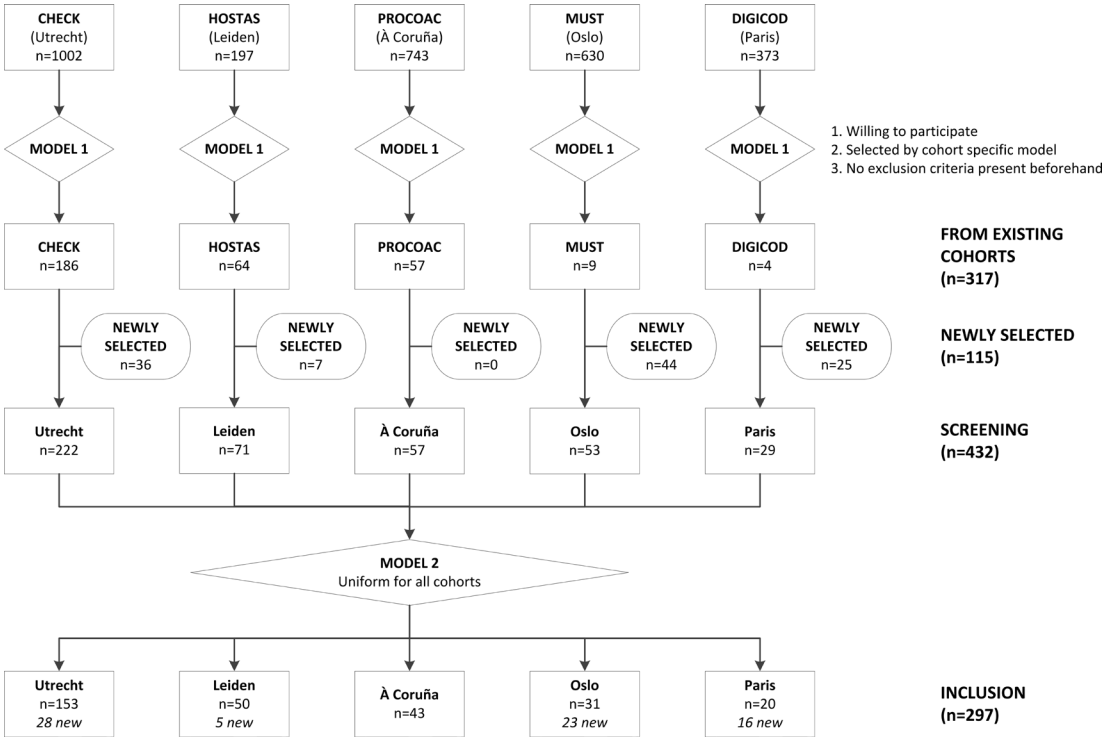

Figure 1 Flow diagram of patient selection for the APPROACH cohort study. APPROACH, Applied Public-Private Research Enabling OsteoArthritis Clinical Headway

- ≥18 years of age.
- Capable of understanding the study.
- Capable of communicating in local language.
- Predominantly tibiofemoral knee OA and satisfying the clinical ACR classification criteria for knee OA:
  - Knee pain.
  - Three or more of the following:
    - >50 years of age.
    - <30 min of morning stiffness.
    - Crepitus on active motion.
    - Bony tenderness.
    - Bony enlargement.
    - No palpable warmth.
- High probability of progression (ranking) based on the algorithm using the following parameters:
  - Reduced version of KOOS questionnaire (pain, stiffness and function).
  - Body Mass Index.
  - Numeric Rating Scale (NRS) pain[19] of index knee at moment of screening visit.
  - NRS pain of index knee in last week before screening visit.
  - Age.
  - Gender.
  - KIDA parameters of the index knee, based on standard weight-bearing radiograph, taken at screening.[18]

Exclusion criteria

- Inability to comply to the protocol.
- Participation in a trial of local therapeutic intervention for index knee OA or potential systemic DMOADs at the time of inclusion, within six months before inclusion, and/or anticipated during two years of follow-up. Participation in non-interventional studies was allowed.
- Surgery of the index knee in the six months before inclusion and/or scheduled or expected surgery of the index knee during follow-up.
- Current pregnancy or planned pregnancy during follow-up (because of imaging).
- Predominantly patellofemoral knee OA.
- Secondary knee OA. For example, due to severe leg deformity (knee varus or valgus >10°), inflammatory joint disease (either autoimmune, infectious or crystal-induced), severe chondrocalcinosis, Paget's disease of the bone, ochronosis, acromegaly, haemochromatosis, Wilson's disease, osteochondritis dissecans, haemophilia.
- Alternative/additional causes of joint pain, for example, rheumatic symptoms due to malignancies, primary osteochondromatosis, osteonecrosis.
- Generalised pain syndrome, for example, fibromyalgia.
- Patients with contraindications for undergoing MRI or CT.
- Previous hip replacement or expected hip replacement within six months.
- Osteosynthesis material near the knee joint.
- Self-reported severe spine OA.
- Current knee prosthesis; in case of surgical replacement of the index or contralateral knee during follow-up, images of that joint will be considered irrelevant and not be obtained. All other acquisitions will be performed as scheduled and patients will remain in the study.

Baseline characteristics of the APPROACH cohort

The baseline characteristics of the APPROACH cohort in total and per centre are shown in table 1.

Despite ranking of all screened patients from the different cohorts in one uniform ML model, baseline characteristics differed between the patients that were included from the different cohorts, representing the characteristics of the original source cohorts.

Investigation schedule

Conventional and novel clinical, imaging, biochemical and kinetic markers of the index knee and other joints were obtained at baseline and will be obtained at 6, 12 and 24 months (table 2). For a detailed description of all parameters see online supplementary file 2.

Parameters for description of OA progression and phenotypes

OA progression and phenotype of the index knee over two years will be described by changes from baseline to the 1-year and/or 2-year visit.

The parameters used to define structural progression will be:

- Radiographic parameters of knee OA severity; JSW and osteophytes using KIDA measurements, Kellgren and Lawrence (KL) grading, and Osteoarthritis Research Society International (OARSI) grading.[18 20 21]
- Quantitative MRI parameters for cartilage including thickness, volume and denuded bone areas in the femorotibial joint.[22]
- Semi quantitative MRI scoring of cartilaginous and non-cartilaginous components including bone marrow oedema, meniscal alteration and synovitis, assessed separately and under a global score.[23]
- Advanced radiographic parameters; bone shape analyses and subchondral bone architecture on standard radiographs and high-resolution CT representing OA related bone and trabecular deformations/adaptations.[24]
- (Bio)markers in blood and urine representing cartilage, bone and synovial matrix turnover and inflammation.

The parameters for pain and function will be:

- KOOS questionnaire.[17]
- Knee Intermittent and Constant OsteoArthritis Pain (ICOAP) questionnaire.[25]
- General pain and function parameters (eg, physical examination index knee).

Parameters for prediction of index knee OA progression and phenotypes

Prediction of OA progression (phenotype specific) will be evaluated using ML taking into account the parameters

Table 1 Baseline characteristics of the APPROACH cohort study Kellgren and Lawrence (grade)

	Total (n=297)	Utrecht (n=153)	Leiden (n=50)	À coruña (n=43)	Oslo (n=31)	Paris (n=20)	P value (ANOVA)
Age (years)	66.5 (7.1)	67.5 (6.5)	65.0 (7.0)	66.1 (6.9)	64.6 (8.9)	66.8 (8.8)	0.106
Female (%)	230 (77)	109 (71)	39 (78)	39 (91)	23 (74)	20 (100)	**0.008**
BMI (kg/m^2)	28.1 (5.3)	27.1 (4.4)	27.4 (5.2)	31.3 (5.9)	28.7 (6.4)	29.3 (6.0)	**<0.001**
KOOS							
Symptoms	69.5 (17.2)	75.2 (15.7)	65.5 (19.9)	61.9 (13.0)	63.7 (16.0)	62.0 (17.7)	**<0.001**
Pain	66.4 (18.8)	73.1 (17.1)	66.8 (19.1)	52.9 (12.7)	56.4 (17.1)	58.9 (19.9)	**<0.001**
Function	69.1 (19.0)	76.6 (16.5)	69.7 (20.9)	54.0 (10.0)	60.5 (17.3)	56.9 (19.3)	**<0.001**
Physical activity	42.9 (26.8)	52.1 (27.2)	38.0 (27.4)	28.5 (11.6)	31.8 (23.6)	33.8 (25.7)	**<0.001**
Quality of life	52.9 (20.7)	60.5 (19.0)	52.7 (18.9)	38.7 (12.5)	45.2 (19.1)	39.7 (27.8)	**<0.001**
NRS pain (0–10)							
Index knee	4.6 (2.7)	3.8 (2.6)	4.3 (2.6)	6.7 (2.0)	5.4 (2.4)	5.7 (2.8)	**<0.001**
KIDA							
Mean JSW index knee (mm)	5.5 (1.0)	5.6 (1.0)	5.4 (1.0)	5.3 (1.1)	5.2 (1.1)	5.3 (0.9)	0.158
Minimum JSW index knee (mm)	2.5 (1.3)	2.7 (1.2)	2.5 (1.3)	2.3 (1.1)	1.8 (1.3)	2.6 (1.3)	**0.008**
KL grade							**0.048**
Grade 0	51 (17%)	36 (24%)	6 (12%)	7 (16%)	0 (0%)	2 (10%)	
Grade 1	90 (30%)	41 (27%)	18 (36%)	14 (33%)	11 (36%)	6 (30%)	
Grade 2	88 (30%)	37 (24%)	14 (28%)	17 (40%)	11 (36%)	9 (55%)	
Grade 3	54 (18%)	30 (20%)	10 (20%)	3 (7%)	9 (29%)	2 (10%)	
Grade 4	10 (3%)	8 (5%)	1 (2%)	1 (2%)	0 (0%)	0 (0%)	

Continuous variables are given as mean values, SD between brackets and categorical variables as total number, percentages between brackets. Differences between sites were evaluated using ANOVA followed by Tukey's post hoc test. Statistically significant p-values are given in bold.
ANOVA, analysis of variance; APPROACH, Applied Public-Private Research enabling OsteoArthritis Clinical Headway; BMI, body mass index; JSW, joint space width; KIDA, knee image digital analysis; KL, Kellgren and Lawrence; KOOS, knee injury and osteoarthritis outcome score; NRS, Numeric Rating Scale.

mentioned above in addition to explorative markers at baseline and, if available, at 6 months:

- Qualitative MRI parameters; T2 relaxation MRI representing cartilage collagen distribution.[26]
- Advanced radiographic imaging parameters; bone shape analysis on MRI representing bone area and shape.[27]
- Motion analysis (GaitSmart[28]).
- Performance based tests (40m self-paced walk test and 30s chair stand-up test).[29]

Covariables

Additionally, to the above-mentioned parameters, the following covariables are available for the ML modelling and analyses:

- Contralateral knee OA.
  - KIDA measurements,[18] KL grading[20] and OARSI grading.[21]
- Hand OA
  - Inflammation of hand joints (HandScan).[30]
  - OA features of hand joints on standard radiographs: KL grading,[20] OARSI scoring[21] and Verbruggen-Veys grading.[31]
  - Functional Index for Hand Osteoarthritis questionnaire.[32]
- Hip OA
  - OA features of the hips: Whole Body Low Dose CT (WBLDCT).[33]
  - Hip disability and Osteoarthritis Outcome Score[34] and hip ICOAP.[25]
- Facet joint OA and intervertebral disc degeneration: WBLDCT.
- OA features of glenohumeral and acromioclavicular joints: WBLDCT.
- General pain and function parameters
  - Short Form 36 questionnaire for quality of life.[35]
  - Pain with concomitant pain medication registration in a custom made 1-month pain diary.
  - Pain NRS of contralateral knee, both hips, both hands and spine.[19]
  - PainDETECT questionnaire used to identify the likelihood of a neuropathic pain component.[36]
  - Motion analysis (GaitSmart) at 24 months.[28]
  - Performance based tests at 24 months.[29]
  - Physical examination of contralateral knee, hips and hands.

Table 2 Investigation schedule of the APPROACH cohort study

	Screening	BL	6 months	12 months	24 months
Medical history	X	X	X	X	X
General physical examination					
Height	X				
Weight	X	X	X	X	X
Waist circumference		X	X	X	X
Blood pressure and pulse rate		X	X	X	X
Joint examination					
ACR criteria for knee OA	X				
Knee		X	X	X	X
Hand		X	X	X	X
Hip		X	X	X	X
Radiography					
Index knee	X		X	X	X
Contralateral knee		X			X
Hands		X			X
CT-scan					
Index knee		X			X
Whole Body Low Dose CT		X			X
MRI-scan of index knee					
Thickness and volume of cartilage, denuded bone area		X	X	X	X
MOAKS assessment		X	X	X	X
T2-mapping		X	X		
Hand scan		X			X
Motion analysis		X	X		X
Performance-based tests					
40-metre self-paced walk test		X	X		X
30 s chair stand-up test		X	X		X
Questionnaires					
KOOS (pain, stiffness and function)	X				
KOOS		X	X	X	X
HOOS		X			X
ICOAP index knee		X	X	X	X
ICOAP hip		X			X
FIHOA		X			X
Pain NRS index knee	X	X	X	X	X
Pain NRS other joints	X	X	X	X	X
PainDETECT		X	X	X	X
SF-36		X	X	X	X
One month pain diary		X	X	X	X
Biological samples					
Serum		X	X	X	X
Plasma		X			
DNA/RNA		X			X
Urine		X	X	X	X

ACR, American College of Rheumatology; APPROACH, Applied Public-Private Research enabling OsteoArthritis Clinical Headway; BL, baseline; FIHOA, Functional Index for Hand Osteoarthritis; HOOS, Hip disability and Osteoarthritis Outcome Score; ICOAP, Intermittent and Constant OsteoArthritis Pain; KOOS, Knee Injury and Osteoarthritis Outcome Score; MOAKS, MRI osteoarthritis knee score; NRS, Numeric Rating Scale; OA, osteoarthritis; SF-36, Short Form 36.

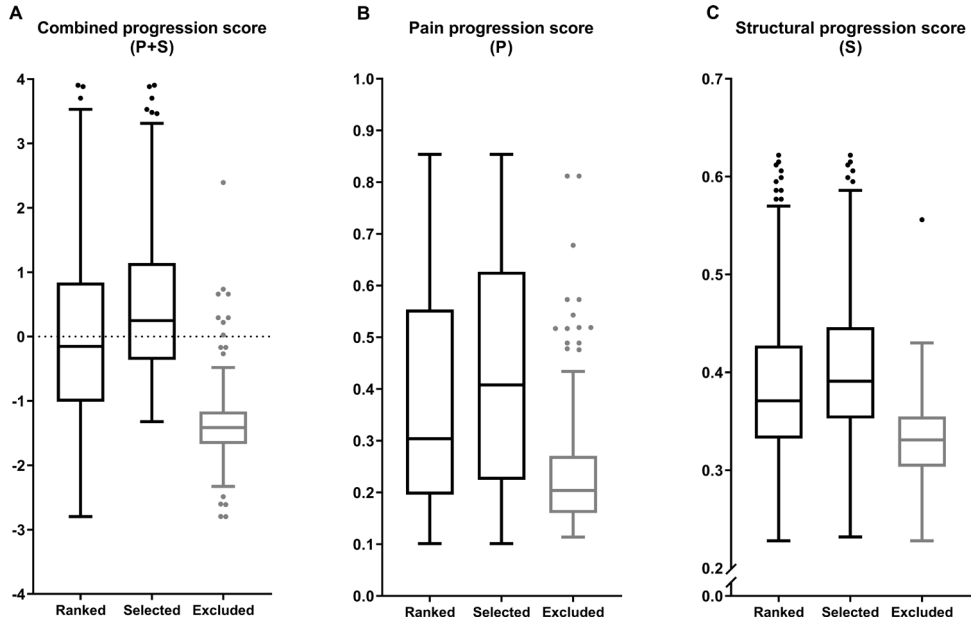

Figure 2 Predicted progression scores of the approach participants. combined (A), pain (B) and structural (C) progression scores (confidence estimates) of the ranked (n=409), selected (n=314) and excluded (n=112) patients. Boxplots represent mean±IQR.

- Optional systemic biochemical (bio)markers
  - Epigenetic, genomic, transcriptomic, proteomic, lipidomic and metabolomic markers (to be defined).
- General clinical data
  - History and type of knee traumatism and surgery.
  - Smoking habits.
  - Menopausal status.
  - Concomitant OA treatment.
- Advanced parameters
  - Bone shape analysis on radiographs of contralateral knee.
  - Subchondral bone analysis on radiographs of contralateral knee.
  - Bone shape analysis of the hip on WBLDCT.

Statistical analysis

Statistical analyses of baseline data for the current manuscript were performed using SPSS Statistics V.25. For evaluation of differences between included and excluded patients t-tests were used. Analysis of variance was used to compare included patients of the five different centres. P values <0.05 were considered as statistically significant.

Future analysis plan: Statistical analyses will be in line with the objectives of the original project. At time of data analysis the best methods to address the aims of APPROACH will be defined as this systems medicine is a fast evolving field. The final analysis plan will be decided on before database lock. In an overview it will comprise:

Validation of the prediction model used in the inclusion process: Model predictions of pain and structural progression will be compared with actual observed progression over 2-year follow-up.

Development and validation of a predictive model for OA progression: Baseline data and/or change over the first six months follow-up will be used to train and test (ML) models for OA progression. External validation of these models will be needed for implementation in practice.

Discovery and prediction of phenotypes/endotypes: The dataset will be explored by use of different statistical approaches to define subgroups with common characteristics. Identified phenotypes/endotypes will be be selected in discussion with clinical experts and described and predicted in enough detail to be of use in practical OA diagnosis and patient selection.

Patient and public involvement statement

A patient council (PC) was instituted to ensure that patients are represented in APPROACH. The PC contributed to the design of the clinical study and with that helped shape the project with particular consideration for the interests of study participants. The PC will maintain close contact with the researchers throughout the project.

Findings to date

Figure 2 describes the probability of progression, as predicted by the second, uniform ML model using the screening visit data of all patients, those who were finally selected and those that were excluded. For mean predicted progression scores (confidence estimates) and results for separate centres see online supplementary file 3.

Table 3 Screening characteristics of the total study population

	Included (n=297)	Excluded (n=109)	P value (t-test)
Age (years)	66.5 (7.1)	68.1 (7.7)	0.061
Female (%)	230 (77)	80 (71)	0.013
BMI (kg/m^2)	28.1 (5.3)	26.4 (4.4)	0.003
Adapted KOOS*			
Stiffness	38.5 (21.6)	24.3 (18.5)	<0.001
Pain	31.3 (19.7)	17.7 (14.6)	<0.001
Function	32.9 (19.1)	19.6 (16.3)	<0.001
Total	33.1 (18.8)	19.6 (15.4)	<0.001
NRS pain last week (0–10)			
Index knee	4.6 (2.8)	2.6 (2.2)	<0.001
Contralateral knee	2.8 (2.5)	1.6 (2.2)	<0.001
KIDA			
Mean JSW index knee (mm)	5.5 (1.0)	5.5 (1.1)	0.700
Minimum JSW index knee (mm)	2.5 (1.3)	2.1 (1.0)	0.001
KL grade (%)		ND	NA
Grade 0	51 (17)		
Grade 1	90 (30)		
Grade 2	88 (30)		
Grade 3	54 (18)		
Grade 4	10 (3)		

Continuous variables are given as mean values, SD between brackets and categorical variables as total number, percentages between brackets. Differences between included and excluded patients were evaluated using t-tests and X^2 test (gender).
*A number of KOOS questions was used, weighted to provide a score for stiffness, pain and function of 0 (most severe) to 100 (no limitations).
BMI, body mass index; JSW, joint space width; KIDA, knee image digital analysis; KL, Kellgren and Lawrence; KOOS, Knee injury and Osteoarthritis Outcome Score; NA, not applicable; ND, not determined; NRS, Numeric Rating Scale.

Included versus excluded patients

Out of the 314 patients, 297 patients attended their baseline visit and were included in the cohort. The remaining 17 patients withdrew after initial selection or could not attend the baseline visit before the deadline. All presented baseline parameters were statistically significant different between included and excluded patients, except for age and mean JSW (table 3).

The inclusion process of the APPROACH cohort is considered successful. The dual assessment with the additional screening visit and a second ML has demonstrated the practical value of the chosen recruitment strategy. Although one might have expected higher probabilities of progression from the selection process, opportunities for further optimisation were limited due to a narrow time window and available corresponding data from the source cohorts. Nevertheless, results show a clear differentiation in baseline data of selected and excluded patients, with a predicted increased progression probability of the selected patients. In two years, the success of the approach, viz. the true progression of these patients will become clear. The predicted probabilities will not be 100%, so we expect sufficient non-progressive patients, those anticipated with the lowest probabilities for progression, that will serve as controls.

Data from the 2-year longitudinal cohort will provide valuable insights into the relevance of conventional and novel clinical, imaging and biochemical markers. Changes of these markers over the first 6 months will likely extend the ability to predict the likelihood for OA progression at 12 and 24 months (either pain, structural or both pain and structural) and distinguish between different OA phenotypes. New markers to identify relevant OA phenotypes/endotypes based on imaging, locomotion and biochemical/omics methods will be developed and validated. This will enable classification of each knee OA patient on a phenotype-specific progression scale. Ultimately, the APPROACH cohort intends to provide a basis for phenotype tailored trials of potential DMOADs, decrease the required number of study subjects and trial duration, and therewith form the basis for personalised/stratified medicine in OA.

Strengths and limitations

The APPROACH cohort is part of a larger consortium, bringing together a highly qualified and multidisciplinary group of stakeholders in the form of a public–private partnership of engaged, knowledgeable and complementary industrial, academics and patient experts. The APPROACH cohort is unique in its selection process, recruiting patients from existing cohorts based on ML models with encouraging results of which the actual utility needs to be demonstrated at the end of the 2-year follow-up. The APPROACH cohort will provide 2-year follow-up data of 297 knee OA patients including conventional and novel, explorative, imaging, biochemical, clinical and demographic (bio)markers according to strict protocols for acquisition and evaluation with the aim to identify phenotypes and develop predictive models for progression of these phenotypes. The relative limited 2-year follow-up allows translation of results to pragmatic trial design in the future. The main limitations of the study are the descriptive phase in which the study is at present and the still limited number of included patients related to the large number of outcome parameters.

Collaboration

Currently, the data is confidential and only accessible for the partners of IMI-APPROACH. After completion of the project, access rights have to be approved by the IMI-APPROACH Steering Committee. More information on the project can be obtained from the website: www.approachproject.eu.

Author affiliations
[1]Rheumatology and Clinical Immunology, UMC Utrecht, Utrecht, The Netherlands
[2]Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
[3]Rheumatology, Leiden Universitair Medisch Centrum, Leiden, The Netherlands
[4]Servicio de Reumatologia, INIBIC-Hospital Universitario A Coruña, A Coruña, Spain
[5]Rheumatology, Diakonhjemmet Hospital, Oslo, Norway
[6]Rheumatology, Sorbonne Université, Paris, France
[7]School of Computing Science, Newcastle University, Newcastle upon Tyne, UK
[8]Merck Serono Research, Merck KGaA, Darmstadt, Germany
[9]Musculoskeletal Research Group, Newcastle University, Newcastle upon Tyne, UK
[10]Biomarkers and Research, Nordic Bioscience, Herlev, Denmark
[11]Regenarative Medicine, State Research Institute Center of Innovative Medicine, Vilnius, Lithuania
[12]GlaxoSmithKline USA, Philadelphia, Pennsylvania, USA
[13]Lygature, Utrecht, The Netherlands
[14]Orthopaedics, Erasmus University Medical Center, Rotterdam, The Netherlands
[15]Institut de Recherches Internationales Servier, Suresnes, France

Contributors WEvS, PMJW, JBa, CHL, JLo, ACB-J, AM, JLa, JBo, HHW, AL, ACAM and FPJGL were responsible for designing the study protocol. WEvS, EMvH, MPJ, MK, ML, FJB, IKH and FB were responsible for data acquisition. EMvH was responsible for data analyses. EMvH was responsible for drafting the manuscript. All authors revised the manuscript thoroughly and aggreed with the final version.

Funding The research leading to these results have received support from the Innovative Medicines Initiative Joint Undertaking under Grant Agreement no 115770, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. See www.imi.europa.eu and www.approachproject.eu

Disclaimer This communication reflects the views of the authors and neither IMI nor the European Union and EFPIA are liable for any use that may be made of the information contained herein.

Competing interests EMvH has nothing to disclose; WEvS reports grants from The Innovative Medicines Initiative Joint Undertaking under Grant Agreement no 115770, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution, during the conduct of the study; MPJ has nothing to disclose; PMJW has nothing to disclose. MK reports grants from IMI-APPROACH, grants from Dutch Arthritis Association, during the conduct of the study; other from GlaxoSmithKline, Pfizer, Merck-Serono, Kiniksa, Abbvie, outside the submitted work; ML reports grants from Inovative Medicines Initative, during the conduct of the study; FJB reports grants from Gebro Pharma, grants from BIOIBERICA, grants from AB Science, grants from Abbvie, grants from Ablynx N.V., grants from Amgen, grants from Archigen Biotech, grants from Boehringer, grants from Bristol-Myers, grants from Celgene Int., grants from Eli Lilly and Company, grants from F. Hoffmann- La Roche, grants from Galapagos, grants from Gedeon, grants from Genentech, grants from Gideal Sciences, NC, grants from Glaxosmithkline, grants from Hospira, grants from INC Research UK, grants from Inventiv Health Clinical, grants from Janssen, grants from Lilly, grants from Nichi-IKO Pharmaceutical, grants from Novartis, grants from ONO Pharma, grants from Pfizer, grants from Pharmaceutical Research, grants from Regeneron, grants from Roche, grants from SA UCB Pharma, grants from Sanofi, grants from TRB Chemedica, grants from UCB Biosciences GMBH, outside the submitted work; In addition, FJB has a patent Molecular block-matching method for gel image analysis issued, a patent Targeting A Specific Receptor On Cells With A Specific Compound For Use In The Treatment And/Or The Prevention Of Osteoarthritis And Rheumatoid Arthritis pending, a patent Genetic markers for osteoarthritis issued, a patent Method for the diagnosis of osteoarthritis issued, a patent Genetic markers for osteoarthritis pending, a patent Method for the diagnosing Arthrosis pending, a patent Method for diagnosing Arthrosis pending, a patent Method for the diagnosis of osteoarthritis pending, and a patent Anti-connexin compounds for use in the prevention and/or treatment of degenerative joint diseases. pending; IKH reports personal fees from AbbVie, grants from Pfizer, outside the submitted work; FB reports personal fees from Boehringer, Bone Therapeutics, Expanscience, Galapagos, Gilead, GSK, Merck Sereno, MSD, Nordic, Novartis, Pfizer, Regulaxis, Roche, Sandoz, Sanofi, Servier, UCB, Peptinov, TRB Chemedica, 4P Pharma, outside the submitted work; CHL reports other from Merck KGaA, during the conduct of the study; JLo has nothing to disclose; ACB-J reports non-financial support from Nordic Bioscience, personal fees from Nordic Bioscience, during the conduct of the study; AM has nothing to disclose; JLa reports personal fees and other from GlaxoSmithKline, outside the submitted work; and Current employee and shareholder of GlaxoSmithKline; JBo reports grants from Innovative Medicines Initiative (IMI-1), during the conduct of the study; and one of Lygature's other project receives part of its funding directly from Merck KgaA. This project is in the field of schistosomiasis and has no relationship whatsoever with the APPROACH project; HHW has nothing to disclose; ACAM has nothing to disclose; FPJGL has nothing to disclose.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not required.

Ethics approval The study is being conducted in compliance with the protocol, Good Clinical Practice (GCP), the Declaration of Helsinki, and the applicable ethical and legal regulatory requirements (for all countries involved), and is registered under clinicaltrials.gov nr: NCT03883568. All participants have received oral and written information and provided written informed consent.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. In order to gain and govern access to the central APPROACH databases, tranSMART and XNAT, access has to be approved by the APPROACH Steering Committee.

ORCID iDs
Eefje M van Helvoort http://orcid.org/0000-0003-3540-4892
Mylène P Jansen http://orcid.org/0000-0003-1929-6350
Marieke Loef http://orcid.org/0000-0003-4301-8350
Francisco J Blanco http://orcid.org/0000-0001-9821-7635

REFERENCES

1 Loeser RF, Goldring SR, Scanzello CR, *et al*. Osteoarthritis: a disease of the joint as an organ. *Arthritis Rheum* 2012;64:1697–707.
2 Bijlsma JWJ, Berenbaum F, Lafeber FPJG. Osteoarthritis: an update with relevance for clinical practice. *Lancet* 2011;377:2115–26.
3 Vos T, Flaxman AD, Naghavi M, *et al*. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the global burden of disease study 2010. *Lancet* 2012;380:2163–96.
4 Cross M, Smith E, Hoy D, *et al*. The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 2014;73:1323–30.
5 Puig-Junoy J, Ruiz Zamora A. Socio-Economic costs of osteoarthritis: a systematic review of cost-of-illness studies. *Semin Arthritis Rheum* 2015;44:531–41.
6 Palazzo C, Nguyen C, Lefevre-Colau M-M, *et al*. Risk factors and burden of osteoarthritis. *Ann Phys Rehabil Med* 2016;59:134–8.
7 Felson D, Niu J, Sack B, *et al*. Progression of osteoarthritis as a state of inertia. *Ann Rheum Dis* 2013;72:924–9.
8 Amur S, LaVange L, Zineh I, *et al*. Biomarker qualification: toward a multiple Stakeholder framework for biomarker development, regulatory acceptance, and utilization. *Clin Pharmacol Ther* 2015;98:34–46.
9 Wesseling J, Boers M, Viergever MA, *et al*. Cohort profile: cohort hip and cohort knee (check) study. *Int J Epidemiol* 2016;45:36–44.
10 Hunter DJ, Nevitt M, Losina E, *et al*. Biomarkers for osteoarthritis: current position and steps towards further validation. *Best Pract Res Clin Rheumatol* 2014;28:61–71.
11 Widera P. A machine learning APPROACH; to recruitment in OA [abstract]. *Osteoarthritis Cartilage* 2019;27:S15.
12 Damman W, Liu R, Kroon FPB, *et al*. Do comorbidities play a role in hand osteoarthritis disease burden? data from the hand osteoarthritis in secondary care cohort. *J Rheumatol* 2017;44:1659–66.
13 Magnusson K, Hagen KB, Østerås N, *et al*. Diabetes is associated with increased hand pain in erosive hand osteoarthritis: data from a population-based study. *Arthritis Care Res* 2015;67:187–95.
14 Oreiro-Villar N, Fernandez-Moreno M, Cortes-Pereira E, *et al*. Metabolic syndrome and knee osteoarthritis. impact on the

prevalence, severity incidence and progression of the disease. *Osteoarthritis Cartilage* 2017;25:S286–7.
15 Ornetti P, Brandt K, Hellio-Le Graverand M-P, *et al*. OARSI-OMERACT definition of relevant radiological progression in hip/knee osteoarthritis. *Osteoarthritis Cartilage* 2009;17:856–63.
16 Altman R, Asch E, Bloch D, *et al*. Development of criteria for the classification and reporting of osteoarthritis. classification of osteoarthritis of the knee. diagnostic and therapeutic criteria Committee of the American rheumatism association. *Arthritis Rheum* 1986;29:1039–49.
17 Roos EM, Roos HP, Lohmander LS, *et al*. Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure. *J Orthop Sports Phys Ther* 1998;28:88–96.
18 Marijnissen ACA, Vincken KL, Vos PAJM, *et al*. Knee images digital analysis (KIDA): a novel method to quantify individual radiographic features of knee osteoarthritis in detail. *Osteoarthritis Cartilage* 2008;16:234–43.
19 Downie WW, Leatham PA, Rhind VM, *et al*. Studies with pain rating scales. *Ann Rheum Dis* 1978;37:378–81.
20 Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis* 1957;16:494–502.
21 Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. *Osteoarthritis Cartilage* 2007;15 Suppl A:A1–56.
22 Wirth W, Eckstein F. A technique for regional analysis of femorotibial cartilage thickness based on quantitative magnetic resonance imaging. *IEEE Trans Med Imaging* 2008;27:737–44.
23 Hunter DJ, Guermazi A, Lo GH, *et al*. Evolution of semi-quantitative whole joint assessment of knee oa: MOAKS (MRI osteoarthritis knee score). *Osteoarthritis Cartilage* 2011;19:990–1002.
24 Zerfass P, Lowitz T, Museyko O, *et al*. An integrated segmentation and analysis approach for QCT of the knee to determine subchondral bone mineral density and texture. *IEEE Trans Biomed Eng* 2012;59:2449–58.
25 Hawker GA, Davis AM, French MR, *et al*. Development and preliminary psychometric testing of a new OA pain measure--an OARSI/OMERACT initiative. *Osteoarthritis Cartilage* 2008;16:409–14.
26 Wirth W, Maschek S, Eckstein F. Sex- and age-dependence of region- and layer-specific knee cartilage composition (spin-spin-relaxation time) in healthy reference subjects. *Ann Anat* 2017;210:1–8.
27 Bowes MA, Vincent GR, Wolstenholme CB, *et al*. A novel method for bone area measurement provides new insights into osteoarthritis and its progression. *Ann Rheum Dis* 2015;74:519–25.
28 Zügner R, Tranberg R, Timperley J, *et al*. Validation of inertial measurement units with optical tracking system in patients operated with total hip arthroplasty. *BMC Musculoskelet Disord* 2019;20:52.
29 Dobson F, Hinman RS, Roos EM, *et al*. OARSI recommended Performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. *Osteoarthritis Cartilage* 2013;21:1042–52.
30 van Onna M, Ten Cate DF, Tsoi KL, *et al*. Assessment of disease activity in patients with rheumatoid arthritis using optical spectral transmission measurements, a non-invasive imaging technique. *Ann Rheum Dis* 2016;75:511–8.
31 Verbruggen G, Veys EM. Erosive and non-erosive hand osteoarthritis. use and limitations of two scoring systems. *Osteoarthritis Cartilage* 2000;8 Suppl A:S45–54.
32 Dreiser RL, Maheu E, Guillou GB, *et al*. Validation of an algofunctional index for osteoarthritis of the hand. *Rev Rhum Engl Ed* 1995;62:43s–53.
33 Gielis W, Foppen W, Nap FJ, *et al*. Whole body low dose CT to assess overall burden of osteoarthritis: development of an atlas and reliability testing of a new scoring system. *Osteoarthritis Cartilage* 2019;27:S320.
34 Nilsdotter AK, Lohmander LS, Klässbo M, *et al*. Hip disability and osteoarthritis outcome score (HOOS)--validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord* 2003;4:10.
35 Ware J, Snow K, *et al*, The Health Institute, New England Medical Center. Sf-36 health survey manual and interpretation guide. *Boston, MA:* 1993.
36 Freynhagen R, Baron R, Gockel U, *et al*. painDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. *Curr Med Res Opin* 2006;22:1911–20.